International Journal of Mass Spectrometry 309 (2012) 1–12

Contents lists available at ScienceDirect

International Journal of Mass Spectrometry

journal homepage: www.elsevier.com/locate/ijms

Comparative phosphoproteome analysis reveals more ERK activation in MDA-MB-231 than in MCF-7

Mohammad Humayun Kabir[a,b], Eui Jin Suh[a,b], Cheolju Lee[a,b,*]

[a] Life Sciences Division, Korea Institute of Science and Technology, Seongbuk, Seoul 136-791, Republic of Korea
[b] University of Science and Technology, Yuseong-gu, Daejeon 305-350, Republic of Korea

ARTICLE INFO

Article history:
Received 23 May 2011
Received in revised form 1 August 2011
Accepted 1 August 2011
Available online 10 August 2011

Keywords:
Phosphoprotein
TiO_2
MCF-7
MDA-MB-231
ERK
Phosphorylation motif

ABSTRACT

Protein phosphorylation forms the foundation of many intra-cellular signaling networks. Enrichment is very important to analyze phosphoproteins due to their low abundance in cells. We used TiO_2 enrichment and LC–MS/MS to identify phosphopeptides from two breast cancer cells, MCF-7 and MDA-MB-231. The dataset generated were searched using Sequest algorithm with 1% false discovery rate. It was observed that 234 phosphoproteins and 280 phosphopeptides were common from a total of 590 and 455 phosphoproteins and 802 and 599 non-redundant phosphopeptides identified from MCF-7 and MDA-MB-231, respectively. Besides, a total of 843 and 578 non-redundant phosphorylation sites were also identified from these two cell lines. ERK1/ERK2 kinase substrate motif (PX[S/T]#P) was identified in MDA-MB-231 but not in MCF-7, showing that this kinase signal transduction pathway is a key factor during breast cancer progression. Activation of ERK1/ERK2 in MDA-MB-231 was further confirmed by Western blot. Quantitative analysis by spectral counting revealed 31 phosphoproteins were statistically in different levels between the two cells lines. This difference is due to the differences in kinase activities as well as protein expression levels. Since breast cancer research makes extensive use of both of these cell lines, our study's findings will be of value in understanding phosphorylation signal of breast cancer.

1. Introduction

Protein phosphorylation is a complex network of signaling and regulatory events that affects virtually every cellular process in many ways such as by increasing or decreasing biological activity of a protein, stabilizing or marking it for breakdown, facilitating or inhibiting movement between sub-cellular compartments, initiating or disrupting protein–protein interactions [1,2]. Often, phosphorylation occurs at low stoichiometry or can occur on a protein with low expression level; both conditions represent a challenge for the detection of phosphorylation site and could be determining factors limiting the number of phosphorylation sites identified and investigated so far [3].

Breast cancer is the most common malignancy among women and a second leading cause of cancer death. There is evidence that phosphorylation is involved in various breast cancer cellular mechanisms. For example, phosphorylation of retinoblastoma protein (pRB) appears to be regulated mainly by Cdk interacting protein 1 (Cip1/p2l) through the suppression of cyclin E and cyclin dependent kinase 2 (Cdk2) in breast cancer [4]. Multiple phosphorylated forms of estrogen receptor (ER)α can be detected in multiple human breast tumor biopsy samples which establish the relevance of investigating the regulation and function of ERα phosphorylation to human breast cancer [5]. Akt (also known as protein kinase B, PKB), a serine/threonine protein kinase, phosphorylation is revealed as prognostic and/or predictive role in breast cancer as well as prostate and non-small cell lung cancer [6]. Increased 3-phosphoinositide-dependent protein kinase-1 (PDK-1) phosphorylation is significantly correlated with invasive breast carcinoma [7]. From these examples it is apparent that phosphorylation plays significant role in breast cancer malignancy. Cell transformation often results from activation of components in signaling pathways that control cell proliferation and differentiation. These pathways are initiated from various cell surface receptors, and may converge on the MAPK cascade, a module consisting of MAP kinase kinase (MEK) and MAPK. Breast cancer is genetically and clinically heterogeneous. This disease encompasses several distinct entities with remarkably different biological characteristics and clinical behaviors. These subtypes of breast cancer are generally diagnosed based upon the presence or absence of three receptors: estrogen receptors (ER), progesterone receptors (PR) and human epidermal growth factor receptor-2 (HER-2). More than 60% of

* Corresponding author at: Life Sciences Division, Korea Institute of Science and Technology, Seongbuk, Seoul 136-791, Republic of Korea. Tel.: +82 2 958 6920; fax: +82 2 958 6919.
E-mail address: clee270@kist.re.kr (C. Lee).

1387-3806/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijms.2011.08.002

the total breast cancer patients are ER-, PR- and HER-2 positive. However, there are approximately 15% of all types of breast cancer in women who do not express ERα, PR, and HER-2 and, are, thus, defined as triple-negative breast cancer. The triple-negative subtype breast cancer lacks the benefits of specific therapies that target these receptors. Today chemotherapy is the only systematic therapy for patients with triple-negative breast cancer [8–10]. Several cell lines are available for breast cancer study. MCF-7 is the most widely used breast cancer cell line to study cancer biology and hormone mechanism of action [11,12]. MDA-MB-231 is also used extensively as a breast cancer cell line. Most importantly, MCF-7 is an ER positive breast cancer cell line and MDA-MB-231 is a triple-negative breast cancer cell line. Besides, these two cell lines are used for a wide variety of comparative studies like-antiestrogenic effect of TCDD (2,3,7,8-tetrachlorodibenzo-*p*-dioxin) or EROD (7-ethoxyresorufin-*O*-deethylase activity) [13], effects of antibiotics on cell cycle and apoptosis [14], induction of apoptosis by oligonol [15], glycosphingolipid composition [16], effect of tocotrienol [17] and differences in proliferative or invasive properties due to proteoglycan synthesis [18]. Many of the cases, these two cell lines are chosen as model breast cancer cell lines based on ER. However, compared to other types of cancer cell lines, breast cancer cell lines are less studied for phosphoproteome analysis. Therefore, we targeted to analyze phosphoproteome from the two breast cancer cell lines.

It is evident from various research reports that these days' phosphoproteome studies draw growing attention. But analysis of phosphoproteins is not simple due to small portion of protein phosphorylation [19]. A number of strategies to overcome these problems have been developed, involving either physically enriching for phosphorylated peptides or phosphorylated proteins [20]. A very meticulous enrichment procedure for phosphorylated peptides based on titanium dioxide has recently been developed. This metal oxide has selective adsorptive properties to organic phosphates [21,22]. The basis of this approach involves a selective interaction of phosphates with porous titania beads by bidentate binding at the TiO_2 surface [23,24]. Later, a more selective enrichment procedure is reported using 2,5-dihydroxybenzoic acid (DHB), lactic acid or glycolic acid as peptide loading buffer in TiO_2 microcolumns to enrich phosphopeptides where the added acid efficiently reduced the binding of non-phosphorylated peptides to TiO_2 retaining high binding affinity for phosphorylated peptides [25–28]. In our work, we have used DHB as a loading buffer for TiO_2 affinity resin to enrich phosphopeptides from two breast cancer cell lines, MCF-7 and MDA-MB-231. Analytical strategy involves cell culture, harvest, SDS-PAGE, in-gel digestion, TiO_2 enrichment and LC–MS/MS analysis. Our objective is to find phosphoproteins existing in different levels between two breast cancer cell lines as well as some novel phosphoproteins and phosphorylation sites which might be a potential resource for future research to study the biological significance of these phosphoproteins and their phosphorylation sites in different signaling pathways related to breast cancer development.

2. Experimental procedures

2.1. Materials

Phosphatase inhibitor calyculin A, okadaic acid and cypermethrin were purchased from Calbiochem (a brand of EMD Biosciences, Inc. La Jolla, CA); sodium fluoride and sodium orthovanadate was from Sigma (Sigma–Aldrich, Inc., St. Louis, MO). Protease inhibitor cocktail was obtained from Roche (Mannheim, Germany). Sequencing-grade modified trypsin was purchased from Promega (Madison, WI). 2,5-Dihydroxybenzoic acid (DHB) was from Sigma. MonoTip® TiO was bought from GL Sciences Inc (Tokyo, Japan). All other chemicals and reagents were of the highest grade commercially available.

2.2. Cell culture, lysis and protein concentration determination

MCF-7 cells were grown in Dulbecco's modified Eagle's medium (DMEM) (Invitrogen, Grand Island, NY) supplemented with 10% fetal bovine serum (FBS; Invitrogen), 100 units/mL penicillin and 100 μg/mL streptomycin in an incubator at 37 °C under 5% CO_2. Similar conditions were followed while culturing MDA-MB-231 cells. In case of MCF-7, cells were grown around 60–70% confluence and for MDA-MB-231, this was around 80–90%. After growing and before harvesting, cells were washed twice with ice cold phosphate buffered saline (PBS). The cells were then resuspended in ice cold lysis buffer containing 35 mM Tris–Cl, 10% glycerol, 0.5% NP-40 and a protease inhibitor mix. In order to preserve protein phosphorylation, several phosphatase inhibitors were also added to a final concentration of 20 nM calyculin A, 200 nM okadaic acid, 4.8 μM cypermethrin, 2 mM vanadate, 10 mM sodium pyrophosphate, 10 mM NaF and 5 mM EDTA [29]. After 10 min incubation on ice, cells were lysed by sonication. Cell debris and nuclei were removed by centrifugation for 15 min at 4 °C using 12,000 rpm. The supernatant was centrifuged one more time for 10 min at the same speed to remove any remaining small particles. Finally, supernatant was collected and store at −80 °C until use. Protein concentration was determined by BCA method.

2.3. Preparative SDS-PAGE and in-gel digestion

Approximately 1 mg of MCF-7 or MDA-MB-231 cell protein was separated by a hand-poured preparative SDS-PAGE (12%) gel (10 cm × 8 cm × 0.10 cm). Electrophoresis was stopped when the buffer front had migrated ∼5 cm into the gel. The gel was stained with Coomassie blue and then excised into 15 regions (∼4 mm × 100 mm). Each gel band was cut into approximately 1-mm cubes, washed with distilled water and destained using 25 mM NH_4HCO_3 in 50% CH_3CN. Gel pieces were dehydrated with 100% CH_3CN and dried in air at room temperature (25 °C). Later, the dried gel pieces were reduced with 10 mM dithiothreitol (DTT) at 56 °C for 1 h and then alkylated with 55 mM iodoacetamide for 1 h in the dark at room temperature (25 °C). After washing with 25 mM NH_4HCO_3 in 50% CH_3CN, gel pieces were again dehydrated with 100% CH_3CN and dried in air at room temperature (25 °C). For trypsin digestion, precooled dry gel pieces were rehydrated with 12.5 ng/μL trypsin (in 25 mM NH_4HCO_3) on ice for 30–45 min. Washing excess trypsin with 25 mM NH_4HCO_3, digestion was carried out at 37 °C for overnight. Digests were extracted once with 25 mM NH_4HCO_3 in 50% CH_3CN, once with 0.2% trifluoroacetic acid (TFA) in 50% CH_3CN and finally once with 0.2% TFA in 70% CH_3CN, dried completely under vacuum, and stored at −20 °C until use.

For whole proteome analysis, the samples were desalted, reconstituted in 0.4% acetic acid, and analyzed by LC–MS/MS. For phosphoproteome analysis, phosphopeptides were enriched by using TiO_2 pipette tip and desalted before LC–MS/MS.

2.4. Enrichment of phosphopeptides

Enrichment was done following manufacturer's instruction with some modifications as described below. The dried tryptic peptide samples were resuspended in 200 μL of 100 mg/mL DHB and 0.1% formic acid in 80% CH_3CN solution. Individual TiO_2 pipette tip was preconditioned with 100% CH_3CN solution and conditioned with 0.2 M phosphate buffer (pH 7.0) by aspirating and dispensing. The pipette tip was then equilibrated with 0.1% formic acid in 50% CH_3CN solution. The peptide sample was adsorbed for around 20

cycles by aspirating and dispensing slowly. The adsorbed peptide in the tip subsequently was rinsed several times with 0.1% formic acid and 0.1 M KCL in 30% CH_3CN solution. Later, phosphopeptides were eluted in 0.2 M phosphate buffer (pH 7.0) and dried under vacuum. Finally, each of the enriched peptide samples was desalted using C_{18} SPE and stored at −20 °C.

2.5. Peptide analysis by LC–MS/MS

Samples were analyzed on a LTQ electrospray ionization linear single-quadruple ion trap mass spectrometer (Thermo, San Jose, CA). Chromatographic separation of peptides was achieved on an Agilent Series 1200 LC system (Agilent Technologies, Waldbronn, Germany), equipped with a 12 cm fused silica emitter with 76 μm inner diameter (BGB Analytik, Böckten, Switzerland), packed in-house with a Magic C18 AQ 5 μm resin (Michrom BioResources, Auburn, CA). Each of the enriched peptide samples or non-enriched samples was dissolved in 0.4% acetic acid prior to loading into chromatography. Peptides were loaded from a cooled (4 °C) Agilent auto sampler and separated with a linear gradient of CH_3CN/water, containing 0.10% formic acid, with a flow rate of 400 nL/min. The gradient was operated from 5 to 40% CH_3CN for first 90 min and 40–80% CH_3CN for next 25 min. The MS survey was scanned from 300 to 2000 m/z followed by three data-dependent MS/MS scans with following options: isolation width, 1.5 m/z; normalized collision energy, 25%; dynamic exclusion duration, 180 s.

2.6. Database search and data process

Using Sequest search engine [30], we searched all MS/MS spectra against International Protein Index human protein sequence database (IPI, versions 3.24, European Bioinformatics Institute, www.ebi.ac.uk/IPI/), common contaminant proteins, and the same human and contaminant protein sequences in reversed orientations. The reverse-database (decoy) was placed after the forward-database (target). Searches were performed with the following parameters: partial-trypsin specificity, peptide mass tolerance of ±3.0 Da, fragment ion mass tolerance of ±0.5 Da, maximum missed cleavages of 2, a static modification of carboxyamidomethylated Cys (+57.0215), and dynamic modifications of phosphorylated Ser, Thr, and Tyr (+79.9799). No dynamic modification was used during the analysis of whole proteome data (non-enriched peptides). All peptide matches were first filtered using ΔCn (normalized difference between XCorr values of the top-ranked candidate peptide and the next candidate with a different amino acid sequence) value of >0.08 and classified according to their charge and tryptic termini. XCorr cutoff filter was then applied to each group of peptides in order to keep false discovery rate (FDR) below 1.0%. It was assumed that true-positive (TP) peptide identifications were exclusively assigned to peptides from target database element and false positive (FP) identifications were equally distributed between target and decoy elements. Therefore, FDR was estimated by doubling the number of decoy database identifications and dividing the result by the number of total peptide identifications in the data set [31].

2.7. Motif analysis

Phosphopeptide sequences were submitted to the Motif-X algorithm (http://motif-x.med.harvard.edu/) [32] for the discovery of phosphorylation motifs present in our data set. Human IPI database was used as background. Sequences were centered on each phosphorylation site and extended to 13 amino acids (±6 residues at both ends of phosphorylation site). Sites which could not be extended due to early occurrence of N or C termini were excluded by the Motif-X algorithm. The significance threshold was set to $P < 10^{-6}$. The minimum number of motif occurrences was set to 5 for both phosphorylated serine and threonine.

2.8. Detection of differential phosphorylation by spectral counting

We used spectral count statistics to measure differential phosphoprotein abundances as well as total protein abundances between MCF-7 and MDA-MB-231 using G-test [33]. The spectral count for each phosphoprotein was corrected according to "Yates' correction for continuity" [34,35]:

$$f_1 = S_1 \pm 0.5 \quad [\text{if } S_1 > S_2 \text{ then}, f_1 = S_1 - 0.5 \text{ or if } S_1 < S_2 \text{ then}, f_1 = S_1 + 0.5]$$

$$f_2 = S_2 \pm 0.5 \quad [\text{if } S_2 > S_1 \text{ then}, f_2 = S_2 - 0.5 \text{ or if } S_2 < S_1 \text{ then}, f_2 = S_2 + 0.5]$$

where S_1 and S_2 denote spectral counts of individual proteins from MCF-7 and MDA-MB-231, respectively. No correction was done if $S_1 = S_2$. The corrected spectral counts were then normalized as:

$$\bar{f}_1 = \frac{f_1}{t_1} \times \frac{t_1 + t_2}{2}; \qquad \bar{f}_2 = \frac{f_2}{t_2} \times \frac{t_1 + t_2}{2}$$

where t_1 and t_2 are total spectral counts obtained from MCF-7 and MDA-MB-231, respectively. Finally, the G-value was calculated as:

$$G = 2\bar{f}_1 \ln\left(\frac{2\bar{f}_1}{\bar{f}_1 + \bar{f}_2}\right) + 2\bar{f}_2 \ln\left(\frac{2\bar{f}_2}{\bar{f}_1 + \bar{f}_2}\right)$$

Proteins with a G-value larger than 3.841 were regarded as differentially expressed with $P < 0.05$ according to χ^2-distribution.

2.9. Western blot analysis

Cells were harvested, washed with PBS, and lysed in the lysis buffer. After 10 min incubation on ice, the cells were sonicated and spun at 12,000 rpm, 4 °C for 10 min to precipitate insoluble debris. Proteins were fractionated by SDS-PAGE, transferred onto a PVDF membrane (Amersham BioScience, Piscataway, NJ), blocked with 5% skim milk in TTBS (20 mM Tris, pH 7.4, 150 mM NaCl, 0.02% sodium azide and 0.05% Tween 20), and incubated with specific antibodies.

Antibodies directed to p44/42 MAPK (ERK1/ERK2) (Cell Signaling, Danvers, MA), phospho-p44/42 MAPK (ERK1/ERK2) (Thr202/Tyr204) (D13.14.4E) XP™ rabbit monoclonal antibody (Cell Signaling, Danvers, MA), MEK-2 (C-16) (SantaCruz Biotechnology Inc., Santa Cruz, CA), phospho-MEK1/2 (Ser217/221) (41G9) rabbit monoclonal antibody (Cell Signaling), α-tubulin and β-actin (SantaCruz Biotechnology Inc.) were used as primary antibodies. Concentrations for blotting antibodies varied according to the manufacturer's recommendations. Blots were washed five times with TTBS buffer (without 0.02% sodium azide) and incubated with horseradish peroxidase-conjugated secondary antibody with 5% skim milk (without 0.02% sodium azide) for 1 h at room temperature, and then developed with a chemiluminescence detection system (Amersham BioScience, Piscataway, NJ).

3. Results and discussion

3.1. Overall strategy

The overall workflow in this study is outlined in Fig. 1. Briefly, 1 mg of protein from MCF-7 and MDA-MB-231 cell was first analyzed separately using a one-dimensional SDS-PAGE. The entire gel

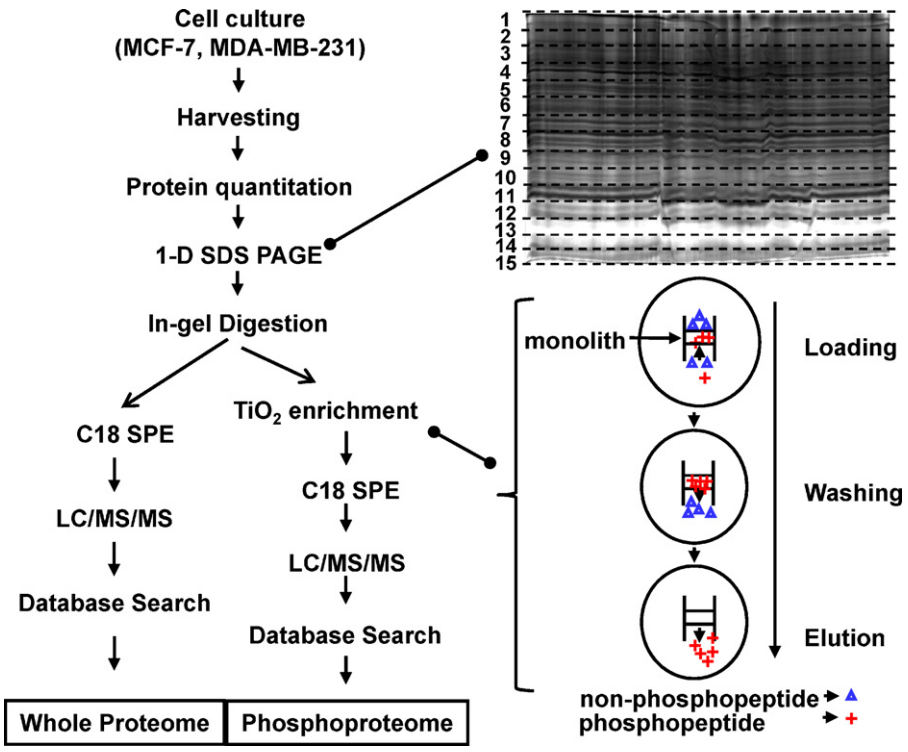

Fig. 1. Flowchart of the strategy used for the phosphoproteome/whole proteome analysis from MCF-7 and MDA-MB-231. Cells were harvested after cell culture. After determining protein concentration, 1 mg of protein was loaded into a 1-D SDS PAGE. The whole gel was sliced into 15 horizontal fractions and trypsin digestion was done for individual fraction. Phosphopeptides were enriched from the digested sample using a TiO_2 pipette tip. The enriched sample was desalted using C_{18} SPE before LC/MS analysis. For whole proteome analysis, samples were desalted without enrichment. MS/MS spectra acquired over 15 runs were searched against the target-decoy human protein database using Sequest. Guided by the decoy hits, criteria using ΔCn, charge state, tryptic termini and XCorr values were used to filter the search results. The false discovery rate was below 1% for the identified phosphopeptides.

was divided into fifteen horizontal gel band fractions. Proteins in each gel band fraction were proteolyzed with trypsin followed by phosphopeptide enrichment by TiO_2 affinity pipette tip. Finally, each sample fraction was desalted by C_{18} SPE and analyzed by LC–MS/MS. For whole proteome profiling, tryptic peptides were directly analyzed without enrichment. The mass spectra acquired over fifteen runs were searched with Sequest against the composite target-decoy human protein database with variable modifications for phosphorylation on Ser, Thr and Tyr. Data were first filtered using ΔCn value of >0.08 and classified according to their charge state and number of tryptic termini. To keep FDR rate below 1%, XCorr cutoff filter was then applied to each group of peptides.

3.2. Distribution of phosphopeptides and phosphoproteins

A total of 1665 and 1230 phosphopeptides from the target database passed our criteria from the two cell lines, MCF-7 and MDA-MB-231, respectively (Supplementary Table S1). The distribution of these phosphopeptides with their charge state is shown in Fig. 2A. In MCF-7, around 52% of the identified phosphopeptides were doubly charged and 48% phosphopeptides were triply charged whereas these values were 39% and 61% in MDA-MB-231. Finally, 802 and 599 non-redundant phosphopeptides were identified from MCF-7 and MDA-MB-231, respectively. These phosphopeptides are classified based on the number of phosphorylation sites on individual peptides (Fig. 2B). Singly phosphorylated peptides occupied around 91% and 97% of the identified peptides from MCF-7 and MDA-MB-231, respectively. An ideal method should enrich singly as well as multiply phosphorylated peptides. TiO_2 enrichment method has a bias toward the enrichment of more singly phosphorylated peptides. Though identifying singly phosphorylated peptides are worthwhile because this increases confidence

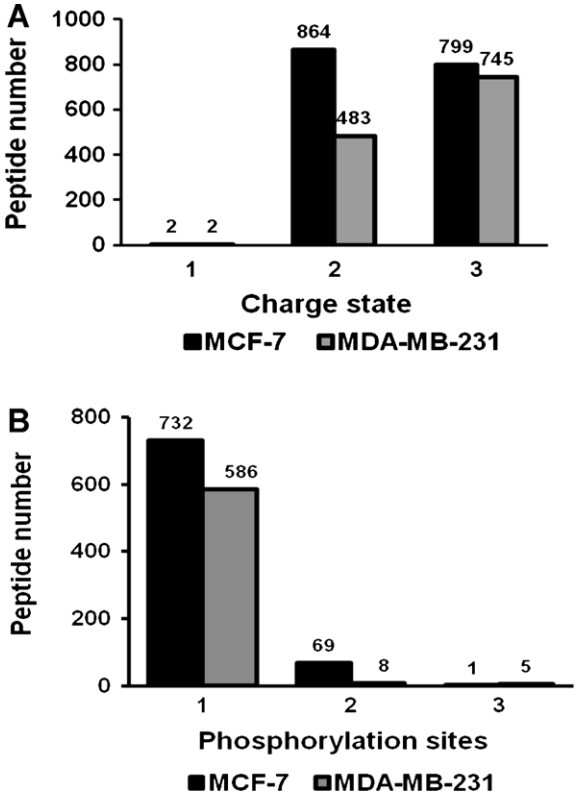

Fig. 2. Distribution of phosphopeptides according to their charge states and the number of phosphorylation sites. (A) Total phosphopeptides were classified according to the charge states. (B) Non-redundant (unique) phosphopeptides were classified according to the number of phosphorylation sites.

Table 1
The number of acidic amino acid residues, prolines, and missed cleavages in the phosphopeptides obtained in our results.

Peptide type	Phosphopeptide	Non-phosphopeptide	PeptideAtlas[a]
Peptide count	1,124	5,561	72,396
Peptide length	21.71	17.65	16.88
Number of D or E	4.56 (21.0%)[b]	3.43 (19.4%)	2.47 (14.6%)
Number of P	2.32 (10.7%)	1.01 (5.7%)	1.12 (6.6%)
Number of missed cleavages	0.68	0.47	0.39

[a] Human PeptideAtlas build 2010 (http://www.peptideatlas.org/).
[b] [Number of residues]/[peptide length] is in parentheses.

in site assignment, it also is a weakness of the TiO_2 enrichment [36–38].

Analyzing the phosphopeptide sequences visually, we observed many peptides with one or more of the following: (1) one to many acidic amino acid residues (aspartic acid or glutamic acid), (2) relatively lengthy sequences, (3) missed cleavages, and (4) proline residues. This visual observation was cross-checked using PeptideAtlas (http://www.peptideatlas.org/) where thoroughly validated tryptic peptides from many diverse proteomic experiments are made available. Peptides from the PeptideAtlas were compared with total phosphorylated and nonphosphorylated peptides obtained from two cell lines in our results for the occurrence rate of acidic amino acid residues, proline residues and of missed cleavages per peptide. The obtained results shown in Table 1 were consistent with our visual observation. Aspartic acid and glutamic acid residue occurred more frequently in our phosphopeptides (21%) than the PeptideAtlas (14.6%). Similar observations were also made concerning peptide length and proline residues. Our analysis correlated well with published reports. For example, phosphopeptide enrichment through TiO_2 or IMAC is biased toward peptides containing negatively charged residues [39] which is a possible reason for many acidic amino acid residues in our results. Moreover, the presence of acidic amino acid residues affects the cleavage of proteases. It is well known that trypsin does not cleave efficiently when an acidic amino acid is located near lysine or arginine residues [39]. Therefore, the identified phosphorylated tryptic peptides are longer, which our study also verified. Likewise, these phosphopeptides are more likely to have missed cleavage sites because of the increased likelihood of aspartic acid and glutamic acid residues occurring adjacent to tryptic cleavage sites [40]. It is

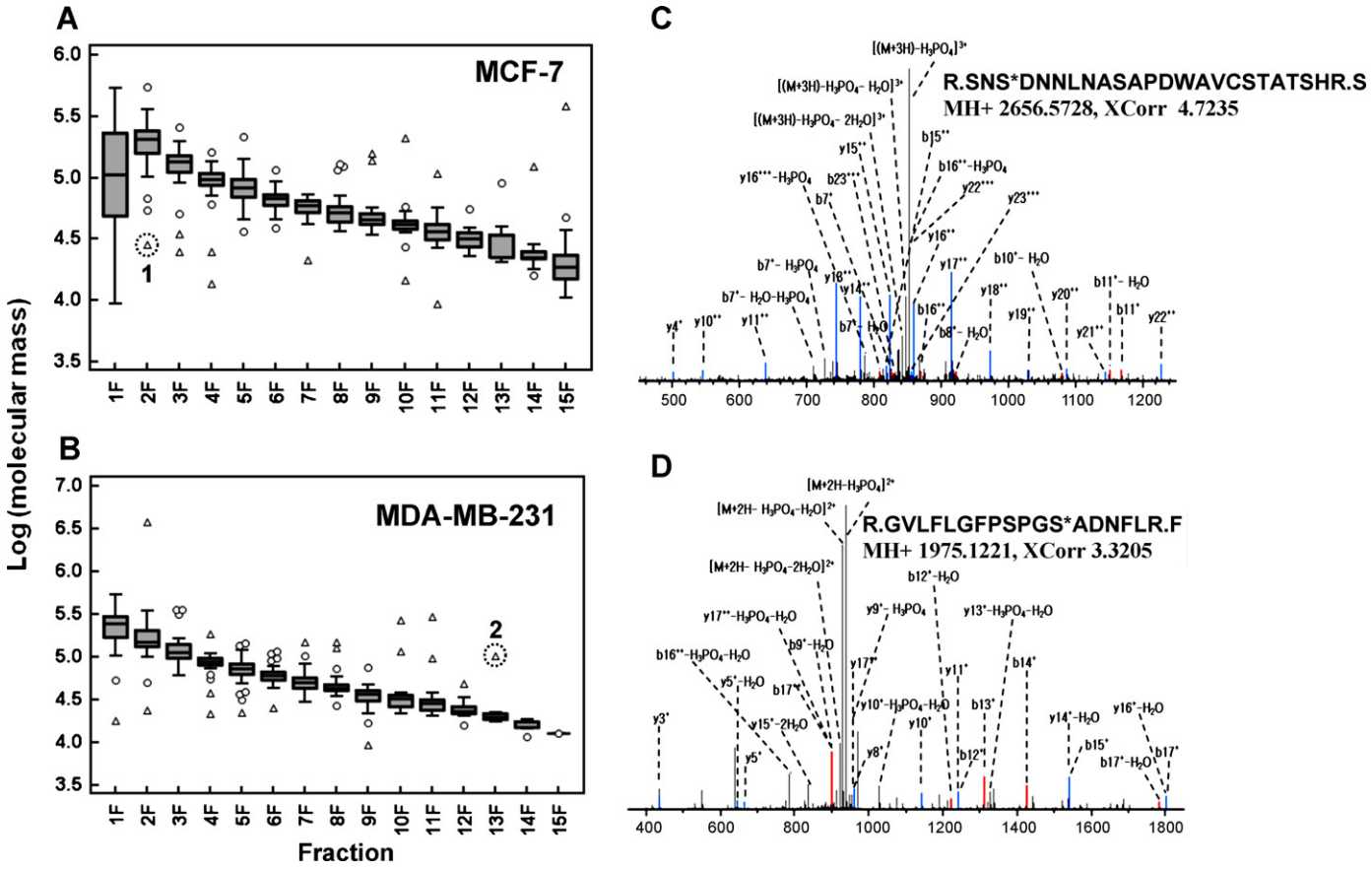

Fig. 3. Box plot for the distribution of molecular weight of the identified phosphoproteins from MCF-7 and MDA-MB-231 in each gel band. (A and B) Individual log molecular mass distribution of the identified proteins from two cell lines. The *X*-axis represents individual gel band fractions and the *Y*-axis represents the log molecular mass. The top, bottom, and line through the middle of the box correspond to the 75th percentile (top quartile), 25th percentile (bottom quartile), and 50th percentile (median), respectively. The whiskers/cap extends from the 10th percentile (bottom decile) and to the 90th percentile (top decile). Circled dot 1 shows a low molecular weight protein identified in the higher molecular weight gel band fraction and circled dot 2 shows a high molecular weight protein identified in the lower molecular weight gel band fraction. (C) MS/MS spectrum of a triply charged peptide from a protein identified in circle 1 (Fig. 3A). (D) MS/MS spectrum of a doubly charged peptide from a protein identified in circle 2 (Fig. 3B).

reported that 50% of all tryptic peptides between 700 and 6000 Da in IPI human protein database contain at least one proline [41]. These proline residues prevent tryptic cleavage when located at the C-terminal side of the lysine or arginine residue. Since more proline residues were found in the enriched peptides, it would result in longer phosphopeptide sequences in our results. Furthermore, we also acquired a nearly equal number of phosphopeptides (MCF-7) or more (MDA-MB-231) with a +3 charge state than a +2 charge state (Fig. 2A). This is also due to trypsin's inefficient cleaving of peptides having acidic amino acid adjacent to lysine or arginine residues, which consequently results in a higher number of positive charge states. A recent report also identified more triply charged phosphopeptides than doubly charged using the metal oxide based enrichment method [38].

A total of 590 proteins from MCF-7 and 455 proteins from MDA-MB-231 were identified. We checked the molecular mass distribution of the identified proteins along the migration distance in SDS-PAGE (Fig. 3). The log molecular mass decreased, as expected, from gel band fraction 1 (the uppermost band in the gel) to gel band fraction 15 (the lowest band of the gel). One notable observation is that individual protein distribution in a few gel band fractions were somewhat distorted by some 'outlier' proteins, e.g., several low molecular mass proteins identified in upper gel band fractions (like dot #1 in Fig. 3A, MCF-7) and several high molecular mass proteins identified in lower gel band fractions (like dot #2 in Fig. 3B, MDA-MB-231). To confirm their identity, MS/MS spectra were checked and excellent spectral quality was observed (Fig. 3C and D). Still the presence of outlier proteins might be due to un-validated false positive identifications (as we used 1% FDR rate), protein identification by shared peptides, different isoforms of a protein, protein degradation, or covalent modification while switching one cellular state to another.

3.3. Analysis of phosphorylation sites

The ΔCn value above 0.08 shows a good compromise to maximize the number of identified peptides and to keep the false positive rate below 1% [31]. Again, phosphopeptides exceeding this value were determined to have localized site(s) of phosphorylation [42]. Beausolil et al. also showed that this parameter directly correlates with the certainty of phosphorylation site assignment and a value greater than 0.15 is required to reach 99.3% certainty [43]. A recent report demonstrated that at a ΔCn value of 0.1 around 9% and at 0.4 around 1% sites were ambiguous [44]. Though we used ΔCn value 0.08 to filter the data, most of the identified phosphopeptides possessed this value above 0.1. Therefore, we can conclude that our identified phosphorylation sites have above a 90% certainty in site assignment. Different lengths of sequences with the same phosphorylation site, just like a sequence ladder, or multiple spectra for the same peptide sequence with same phosphorylation site were observed, which also assured increased confidence in site assignment. Moreover, of the identified phosphorylation sites from the two cell lines, 80.5% and 76.6% had previously been reported by other investigators. This is also a good indication of the reliability of our identified phosphorylated sites.

A total of 843 and 578 non-redundant phosphorylation sites (serine, threonine and tyrosine) were identified from MCF-7 and MDA-MB-231, respectively (Fig. 4). It was observed that the most phosphorylation sites were occupied by serine (85.05% and 85.47%) and the least by tyrosine residue. The ratio of amino acids serine, threonine and tyrosine were 42.18: 6.41: 1 for MCF-7 and 35.29: 5: 1 for MDA-MB-231. Olsen et al. obtained a ratio of 48:6:1 from Hela cells [45], Bodenmiller et al. attained a ratio of 26:6:1 from *Drosophila* Kc167 cell line [44] and Zahedi et al. obtained a ratio of 35:6:2 from resting human platelets [46] for these amino acids. Our

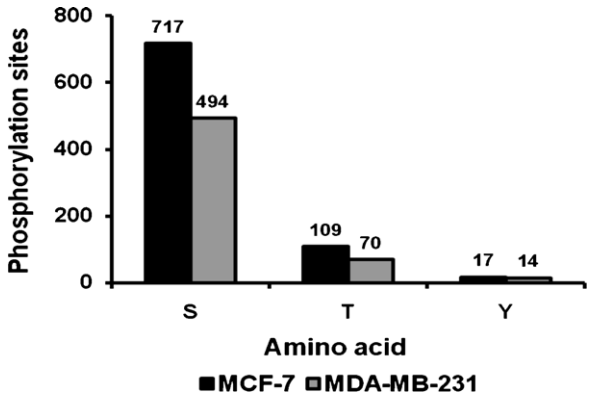

Fig. 4. Distribution of phosphorylation sites occupied by individual amino acid residues (S, T, and Y).

results also exhibit similar findings of these previously performed studies.

3.4. Localization and functional categorization of identified proteins

According to the gene ontology obtained from UniProt database (http://www.uniprot.org), 36% and 39% identified phosphoproteins are located in the nucleus, 26% and 25% are of cytoplasmic origin, while only 10% and 9% are found to be in the membrane of MCF-7 and MDA-MB-231, respectively. But UniProt did not provide any localization data for 16% and 15% phosphoproteins for the two cell lines (Fig. 5A). A closer look at the function revealed that 25% and 23% of the identified proteins are involved in the transcriptional

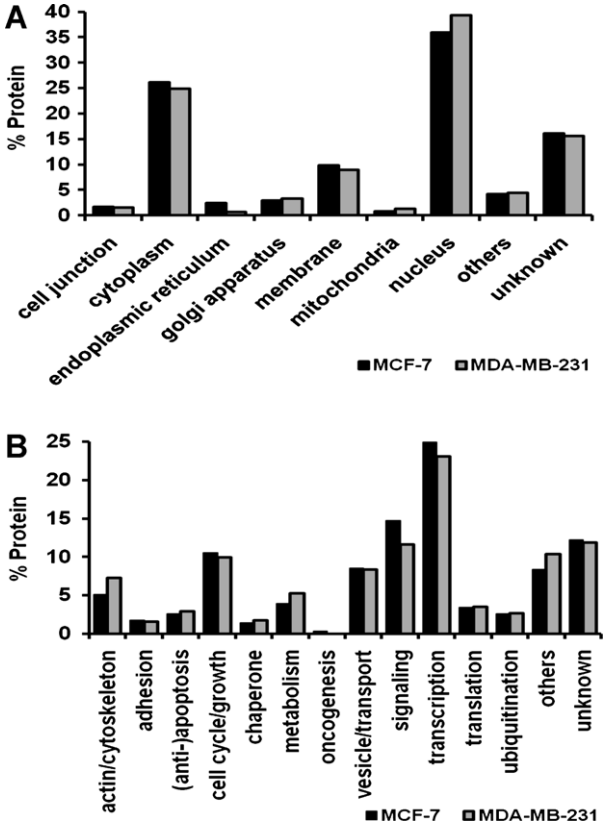

Fig. 5. Classification of the identified phosphoproteins. (A) Phosphoproteins were classified based on their localization. (B) Phosphoproteins were classified based on their involvement in biological processes.

Table 2
List of phosphorylation motifs identified in MCF-7 and MDA-MB-231.

#	Motif[a]	Literature	Score	Phospho dataset (matches/size)		Background data set (matches/size)	
MCF-7							
1.	. . .sP. . .	Proline-directed substrate motif	16.00	280	696	84,335	1,094,911
2.	. . .sD.E. . .	Casein kinase II substrate motif	32.00	80	416	4636	1,010,576
3.	. . .sE.E.E.	Casein kinase II substrate motif	38.11	19	336	883	1,005,940
4.	. . .sDE. . .	?	24.79	17	317	2535	1,005,057
5.	. . .R. . .s. . .	Calmodulin-dependent protein kinase II substrate motif	16.00	62	300	60,451	1,002,522
6.	. . .sE.E. . .	Casein kinase II substrate motif	25.23	22	238	5042	942,071
7.	. . .sDDD. . .	Casein kinase II substrate motif	29.80	11	216	265	937,029
8.	. . .s. . .D. . .	?	11.17	36	205	41,335	936,764
9.	. . .s.E. . .	Golgi-Casein kinase substrate motif	9.12	33	169	51,791	895,429
10.	. . .s.D. . .	Calmodulin-dependent protein kinase II substrate motif Or, PKA kinase substrate motif	7.53	24	136	40,142	843,638
11.	. . .tPP. . .	?	22.57	16	100	4213	687,043
12.	. . .tP. . .	Pro-directed substrate motif	11.77	27	84	44,417	682,830
MDA-MB-231							
1.	. . .P.sP. . .	ERK1/ERK2 kinase substrate motif	23.07	50	471	10,017	1,094,911
2.	. . .sP.K. . .	Cyclin-dependent kinase substrate motif	22.18	21	421	3102	1,084,894
3.	. . .sP. . .	Proline-directed substrate motif	16.00	123	400	71,216	1,081,792
4.	. . .sD.E. . .	Casein kinase II substrate motif	32.00	49	277	4636	1,010,576
5.	. . .R. . .s. . .	Calmodulin-dependent protein kinase II substrate motif Or, PKA kinase substrate motif	16.00	55	228	60,625	1,005,940
6.	. . .sDDD. . .	Casein kinase II substrate motif	30.37	10	173	265	945,315
7.	. . .sE.E. . .	Casein kinase II substrate motif	23.11	21	163	5883	945,050
8.	. . .sD.D. . .	?	18.55	13	142	2730	939,167
9.	. . .D.s. . .	Casein kinase I substrate motif	6.86	22	129	43,537	936,437
10.	. . .tP. . .	Proline-directed substrate motif	16.00	32	63	48,630	687,043

[a] One letter code (in lowercase) represents phosphorylation site. The period represents any amino acid.

process, 15% and 12% are occupied in the signaling process. A total of 11% and 10% proteins are engaged in the cell cycle/growth and 8% proteins are involved in the vesicle/transport from both the cell lines (Fig. 5B). Hence, the two cell lines have similarity with their localization and functional classification of the identified phosphoproteins.

3.5. Analysis of phosphorylation motif

Logo-like representations were created to graphically display each identified phosphorylation motifs (Fig. 6 and Table 2). These logos include information about not only the residues strictly discovered to be part of the motif but also the frequencies of all additional adjacent amino acids. As expected, fewer motifs were found for phosphothreonine (and none for phosphotyrosine) because the total number of phosphorylation events identified on these residues was significantly less than that of serine. However, we have identified twelve (10 serine and 2 threonine) motifs from MCF-7 (Fig. 6A) and ten (9 serine and 1 threonine) motifs from MDA-MB-231 (Fig. 6B), of which five were commonly identified in both cell lines. The vast majority of phosphorylation sites in the data contained acidic ([S/T]#[E/D]), proline-directed ([S/T]#P) and basic (RXX[S/T]#) motifs, in order of abundance. As phosphopeptide enrichment protocols employing TiO_2 and IMAC are biased toward peptides containing negatively charged residues, this bias implies that it is possible that certain phosphorylation motifs are overrepresented in the enriched phosphopeptide data set [39]. Certain motifs are commonly associated with specific kinases as shown in Table 2. But we also have identified significant populations of phosphopeptides containing motifs like [S]#DE, [S]#XXXD [T]#PP, [S]#DXD and that have not yet been associated with a particular kinase.

One of the most noticeable observations is that a vital motif identified in MDA-MB-231 but not in MCF-7 is known as ERK1/ERK2 kinase substrate motif (PX[S/T]#P). In the consensus sequence substrate, proline residues are important for peptide phosphorylation. Proline residue should be located adjacent to the C-terminus and must be at least one amino acid distant from the N-terminus of the site of phosphorylation [47]. This sequence substrate is the target for ERK1/ERK2 serine/threonine kinases (MAPK) that are present in all cell types and serves to link extracellular stimuli to cellular events involved in proliferation and differentiation, including the cell cycle, generation of phospholipid messengers, transcription, and translation [48]. ERK1/ERK2 signal transduction pathway is a key factor during human breast tumorigenesis and breast cancer progression [49]. ER-negative and invasive MDA-MB-231 have constitutively higher ERK1/ERK2 (MAPK) compared to the ER-positive and non-invasive MCF-7 human breast cancer cells. In MCF-7 cells, transforming growth factor (TGF)α stimulation induces only transient MAPK activation, leading to a transient increase in cell migration. However, in MDA-MB-231 cells, TGFα stimulation induces sustained MAPK activation, which correlates with enhanced cell motility.

3.6. Quantitative analysis of phosphoproteins between MCF-7 and MDA-MB-231

Phosphoprotein, phosphopeptide, and phosphorylation site overlap was examined for the two cell lines. It was observed that out of 590 and 455 phosphoproteins identified from MCF-7 and MDA-MB-231, only 234 proteins were common in both cell lines

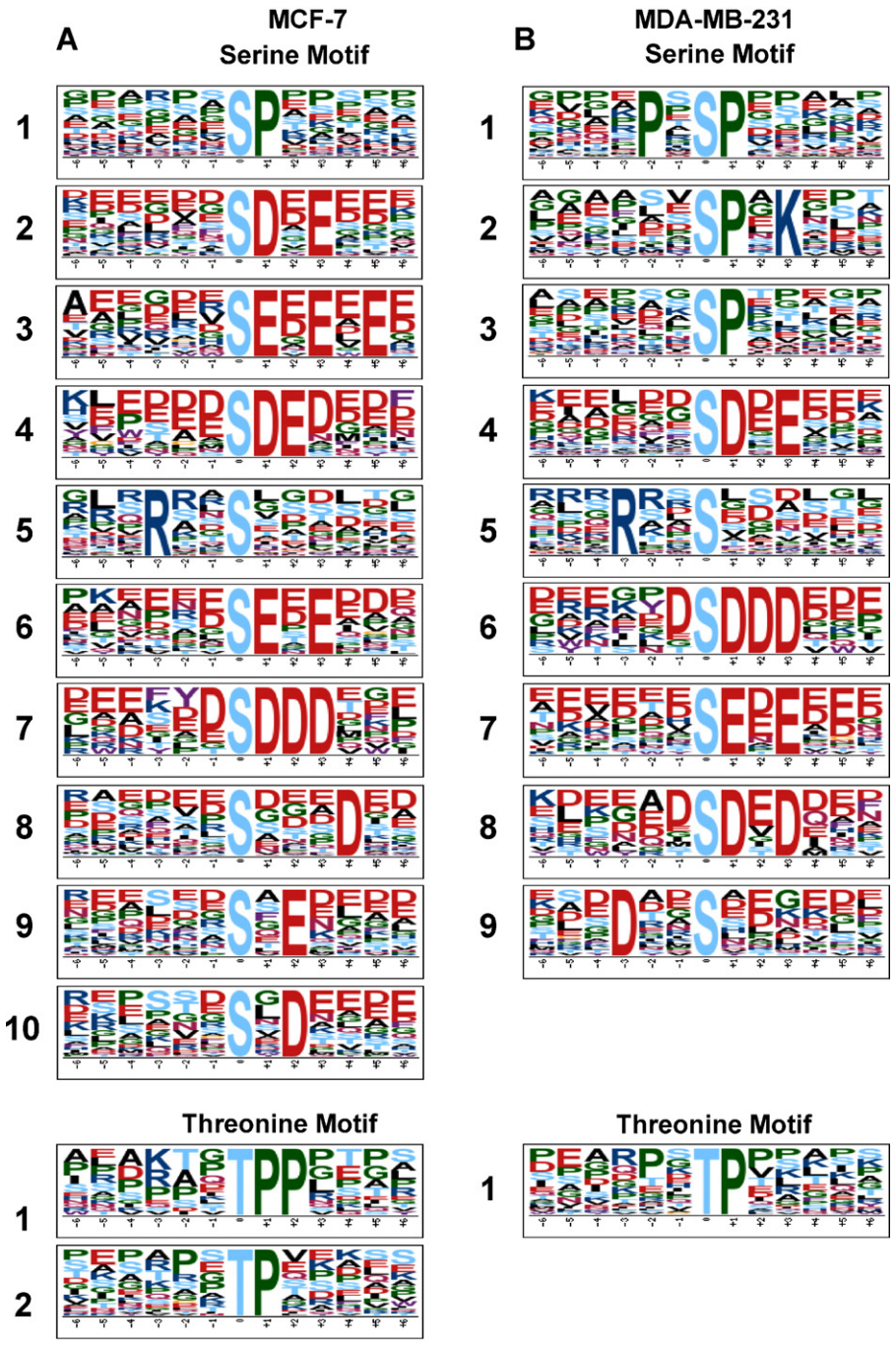

Fig. 6. Sequence logos of the motifs identified from the phosphorylation data set. (A) Motifs identified from MCF-7 and (B) motifs identified from MDA-MB-231 cell line. Site of phosphorylation is shown in blue. The frequency of each residue present in each dataset is proportional to its height. Five (numbers 1, 2, 5, 6 and 7) phosphoserine and 1 phosphothreonine (number 2) motifs identified from MCF-7 are similar with five (numbers 3, 4, 5, 7 and 6) phosphoserine and 1 phosphothreonine (number 1) motifs identified from MDA-MB-231. Acidic amino acids are shown in red and basic amino acids are shown in indigo.

(Fig. 7A). A total of 377 and 341 non-redundant phosphopeptides belonged to these overlapped 234 MCF-7 and MDA-MB-231 proteins of which 280 peptides are shared by both cell lines (Fig. 7B). Besides, 262 phosphorylation sites were common from 843 and 578 non-redundant phosphorylation sites from these two cell lines (Fig. 7C). We also checked the number of proteins identified based on two or more phosphopeptides and their overlapping between the two cell lines. We found 284 and 208 phosphoproteins (having two or more phosphopeptides) of which 110 proteins were shared. We know that phosphorylation status changes with different cellular states. Again, the number of increased missed cleavage sites present in the phosphopeptides due to acidic amino acids (D, E) or proline residues may also contribute to the decreased overlap between the cell lines. Breast cancer cell lines, especially MCF-7 and MDA-MB-231, show very heterogeneous gene expression pattern [50], suggesting that the small number of overlapped phosphorylation sites between the two cell lines may be due to their difference in gene/protein expression pattern. The phosphorylation sites identified from the two cell lines were assessed with the aid of two public databases: PhosphoSite (http://www.phosphosite.org) [51] and UniProt, which have collected comprehensive information on in vivo protein

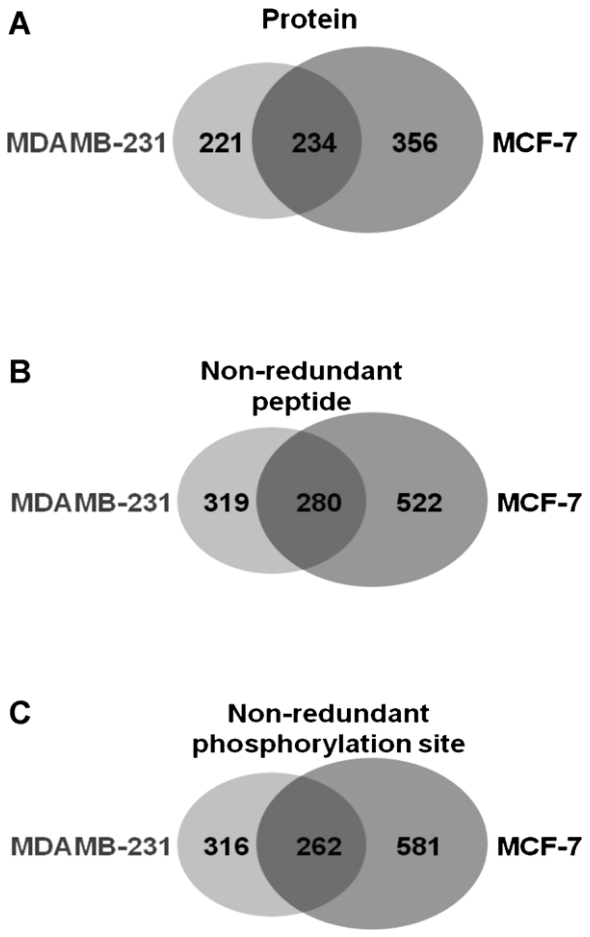

Fig. 7. Comparison of (A) protein and (B) phosphopeptide and (C) phosphorylation site overlap between MCF-7 and MDA-MB-231.

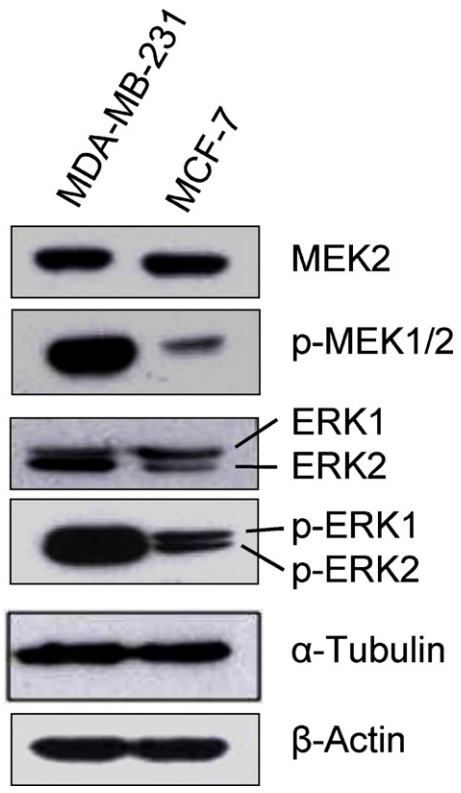

Fig. 8. Upregulation of ERK signaling pathway in MDA-MB-231 compared to MCF-7. The levels of ERK1, ERK2, phospho-ERK1, phospho-ERK2, MEK2, and phospho-MEK1/2 were measured by Western blot. β-Actin and α-tubulin were used as controls of quantification.

phosphorylation and references. Of the phosphorylation sites from these two cell lines, 80.5% (679/843) and 76.6% (443/578) were identified as already known (reported) sites. Approximately 19.5% (164/843) phosphorylation sites and 9.5% phosphoproteins (56/590) identified from MCF-7 cell appeared to be novel; whereas 23.4% (135/578) phosphorylation sites and 9% phosphoproteins (40/455) identified from MDA-MB-231 cell appeared to be novel. A summary for the putative phosphoproteins and their phosphorylation sites is shown in Supplementary Table S2 and the details are listed in Supplementary Table S3.

The difference in phosphoprotein abundance was quantitatively analyzed by spectral counting [33]. The analysis revealed that 31 phosphoproteins existed in statistically different levels between two cells of which 17 were from MCF-7 and 14 were from MDA-MB-231 (Table 3). These abundant proteins are mainly localized in the cytoplasm and nucleus in both of the cell lines but proteins from MDA-MB-231 are more involved in cellular processes like actin/cytoskeleton, chaperon, metabolism, and transcription than in MCF-7 cells. This is one indication of MDA-MB-231's higher invasiveness in nature as a breast cancer cell line than MCF-7 cells. To clarify whether the differences in protein abundances reflect the differences in kinase activities or in protein expression levels between the two cell lines, we also determined the relative protein levels of these proteins through quantitative proteome analysis without enrichment (i.e., whole proteome analysis). *G*-Values and the spectra obtained from total proteome analysis were compared with the *G*-values and the spectra obtained through enrichment step (Table 3). *G*-Value and the number of spectra for several abundant phosphoproteins like calnexin, la-related protein 1, catenin alpha-1, UV excision repair protein RAD23 homolog B, neuroblast differentiation-associated protein AHNAK and large proline-rich protein BAT3 (*G*-value marked with † in Table 3) shows that their abundance is due to the kinase activities in the cell. *G*-Value and the number of spectra for the few abundant phosphoproteins like prostaglandin E synthase 3, serine/arginine repetitive matrix protein 2 and myosin-9 (marked with ‡ in Table 3) show that their abundance is due to the gene/protein expression in the cell. *G*-Values of other abundant proteins do not provide any direct conclusion.

As previously stated, the ERK1/ERK2 kinase substrate motif was enriched in MDA-MB-231 but not in MCF-7 (Fig. 6 and Table 2). To validate the mass spectrometric data, Western blot analysis of ERK and its upstream activator kinase, MEK, was performed. Expressions of ERK1/ERK2 and MEK2 were nearly the same between MCF-7 and MDA-MB-231. However, the levels of phosphorylated forms in the two cell lines differed. In MDA-MB-231, the levels of phospho-ERK1/ERK2 and phospho-MEK1/2 were higher than MCF-7, which shows consistency with the mass spectrometric data (Fig. 8). Besides, there are reports that also revealed the higher ERK1/ERK2 activity in MDA-MB-231 than in MCF-7. For example, Zhou et al. referred a three-fold elevated activity in MDA-MB-231, at the basal level from the nuclear extracts of these two cell lines, than in MCF-7 after the cells grown in complete medium [52]. Another report also confirmed this higher ERK phosphorylation (ERK activity) in non-induced MDA-MB-231 compared to MCF-7 [53].

The phosphoprotein ERK-1 is involved in both the initiation and regulation of meiosis, mitosis, and postmitotic functions in differentiated cells by phosphorylating a number of transcription factors such as ELK-1. Tyr-204, one of the dually phosphorylated residues-Thr-202 and Tyr-204 that activate the enzyme, is identified in our results [54]. Calnexin is an abundant integral

Table 3
Phosphoproteins existing in different levels between MCF-7 and MDA-MB-231.

IPI number	Uniptot ID	Protein name	Number of spectra (phospho)		G-Value (phospho)	Number of spectra (tryptic)		G-Value (tryptic)
			MCF-7	MDA		MCF-7	MDA	
Abundant in MCF-7								
IPI00004560	O15075	Serine/threonine-protein kinase DCLK1; Doublecortin-like and CAM kinase-like 1	13	0	11.24	2	0	0.51
IPI00006196	Q14980	Nuclear mitotic apparatus protein 1; NuMA protein; SP-H antigen	7	0	4.75	7	2	1.79
IPI00007334	Q9UKV3	Apoptotic chromatin condensation inducer in the nucleus; Acinus	10	1	4.57	3	0	1.43
IPI00013788	O43719	HIV Tat-specific factor 1; Tat-SF1	22	3	10.08	–	–	–
IPI00015029	Q15185	Prostaglandin E synthase 3; Hsp90 co-chaperone; Progesterone receptor complex p23	55	12	18.09	16	4	6.27(‡)
IPI00020984	P27824	Calnexin	13	0	11.24(†)	70	81	0.80
IPI00021695	P20020	Plasma membrane calcium-transporting ATPase 1	8	0	5.80	2	0	0.51
IPI00029778	Q12888	Tumor suppressor p53-binding protein 1; 53BP1	8	0	5.80	–	–	–
IPI00185919	Q6PKG0	La-related protein 1; La ribonucleoprotein domain family member 1	12	2	4.07(†)	3	7	0.95
IPI00215948	P35221	Catenin alpha-1; Cadherin-associated protein; Alpha E-catenin; NY-REN-13 antigen	13	1	7.29(†)	16	5	4.84
IPI00240812	Q5R3S3	Sister chromatid cohesion protein PDS5 homolog B	8	0	5.80(†)	3	2	0.00
IPI00304589	Q9C0C2	182 kDa tankyrase-1-binding protein	13	0	11.24	–	–	–
IPI00328293	Q8IYB3	Serine/arginine repetitive matrix protein 1; SRm160	15	3	4.51	2	0	0.51
IPI00334775	Q9NTK6	Heat shock protein HSP 90-beta; HSP 90; HSP 84	16	3	5.24	220	136	18.50
IPI00386170	Q8NEN5	Sperm-specific antigen 2; cleavage signal-1 protein; CS-1; Ki-ras-induced actin-interacting protein	7	0	4.75	3	0	1.43
IPI00464999	Q6AI08	HEAT repeat-containing protein 6; amplified in breast cancer protein 1	15	0	13.48	6	0	4.82
IPI00782992	Q9UQ35	Serine/arginine repetitive matrix protein 2; 300 kDa nuclear matrix antigen	23	4	8.71	4	1	0.80(‡)
Abundant in MDA-MB-231								
IPI00006038	P52948	Nuclear pore complex protein Nup98-Nup96 [Precursor]	0	5	4.87	–	–	–
IPI00008223	P54727	UV excision repair protein RAD23 homolog B	0	8	9.33(†)	7	0	6.04
IPI00014177	Q15019	Septin-2; neural precursor cell expressed developmentally down-regulated protein 5; NEDD-5	4	11	4.60(†)	7	4	0.34
IPI00019502	P35579	Myosin-9	0	22	31.27	46	213	117.19(‡)
IPI00021812	Q09666	Neuroblast differentiation-associated protein AHNAK; Desmoyokin	5	15	7.33(†)	11	12	0.00
IPI00023048	P29692	Elongation factor 1-delta; EF-1-delta; antigen NY-CO-4	25	66	32.16	6	12	1.47
IPI00032206	P50479	PDZ and LIM domain protein 4	0	5	4.87	–	–	–
IPI00178440	P24534	Elongation factor 1-beta; EF-1-beta	13	37	18.94	21	27	0.59
IPI00220740	P06748	Nucleophosmin; NPM; nucleolar phosphoprotein B23; numatrin; nucleolar protein NO38	35	67	21.29	45	72	6.16
IPI00297160	Q86T72	CD44 antigen [Precursor]	0	5	4.87	–	–	–
IPI00398779	Q6S376	Plectin-1	0	5	4.87	–	–	–
IPI00465128	P46379	Large proline-rich protein BAT3; HLA-B-associated transcript 3; Protein G3	5	12	4.36(†)	13	5	2.71
IPI00480085	Q9UQ16	Dynamin-3; dynamin, testicular; T-dynamin	4	13	6.67	–	–	–
IPI00556645	P85037	Forkhead box protein K1	0	5	4.87	–	–	–

Proteins with $G > 3.841$ are counted as differentially expressed with $P < 0.05$ (Ref. [33]).
† shows protein abundance due to kinase activities and ‡ shows protein abundance due to gene/protein expression.

membrane phosphoprotein of the endoplasmic reticulum (ER) of eukaryotic cells. Several kinases have been shown to participate in calnexin phosphorylation. Among them, casein kinase 2 (CK2) was found to be responsible for the phosphorylation of Ser-534 and Ser-544 on human calnexin and ERK1 for Ser-563 [55,56]. Among these sites, a CK2 phosphorylation site, Ser-534 was identified from MCF-7 in our results. On the contrary, no ERK1 phosphorylation site was found. Elongation factor 1 (EF-1) beta is highly phosphorylated by

casein kinase 2 at Serine 106. The beta subunit of eukaryotic EF-1 catalyzes the GDP/GTP exchange activity on EF-1 alpha [57]. Tumor suppressor p53-binding protein 1 (TP53BP1) is phosphorylated on multiple residues in response to different types of DNA damage and is regulated by ATR in response to UV-induced DNA damage [58]. Serine-125 of nucleophosmin (NPM) is indicated to be the major phosphorylation site at the G2 phase. Its over-expression is reported in many types of major human solid tumors including tumors of the colon, liver, stomach, ovary, and prostate [59]. These are some examples of the abundant phosphoproteins and their phosphorylation sites, identified in our result.

4. Conclusions

In this study we have used a modified TiO_2 method to enrich phosphopeptides from two breast cancer cell lines, MCF-7 and MDA-MB-231. Many cell lines (e.g., Hela cell) are commonly used to identify phosphopeptides or phosphorylation sites. But there is only one report so far, to our knowledge, of identifying phosphopeptides from MCF-7 [60] where only a small number of phosphopeptides and phosphorylation sites were identified using a different enrichment method. There is no report known of using MDA-MB-231. Therefore, this study represents the first dataset for the phosphoproteome of MDA-MB-231 cells and the largest phosphoproteome dataset of MCF-7 cells. Since MCF-7 and MDA-MB-231 cell lines are extensively used in breast cancer research, the findings of this investigation will be of value to the breast cancer research community. Exploring phosphoproteins, phosphopeptides and phosphorylation sites, we observed that 234 proteins, 280 phosphopeptides, and 262 phosphorylation sites were common from a total of 590 and 455 phosphoproteins, 802 and 599 non-redundant phosphopeptides, and 843 and 578 phosphorylation sites identified from MCF-7 and MDA-MB-231, respectively. Several motifs including ERK1/ERK2 kinase substrate motif (PX[S/T]#P) were differentiated from the two cell lines. This motif is identified in MDA-MB-231 but not in MCF-7. ERK1/ERK2 signal transduction pathway is a key factor during breast tumorigenesis and breast cancer progression. Quantitative analysis by spectral counting revealed the most prominent features differentiating the two cell lines. We also identified a significant fraction of novel phosphoproteins and novel phosphorylation sites from already known phosphoproteins, which may provide researchers with the opportunity to find different phosphorylation pathways as well as expand the phosphorylation database resource. Finally, the availability of specific antibodies against these phosphorylation sites is also likely to provide an insight into the role of these proteins in different human diseases including cancer.

Acknowledgements

We are grateful to Daniel Schwartz of Harvard Medical School and Bingwen Lu of The Scripps Research Institute for their fruitful comments and discussions while identifying phosphomotifs and protein abundance by spectral counting. We also thank Micheal James Binning for providing his valuable times reading the manuscript thoroughly for correcting the English.

This study was supported by a grant from the Functional Proteomics Center of the 21st Century Frontier R&D Program funded by the Korean Ministry of Education, Science and Technology, Republic of Korea.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ijms.2011.08.002.

References

[1] M. de Graauw, P. Hensbergen, B. van de Water, Phospho-proteomic analysis of cellular signaling, Electrophoresis 27 (2006) 2676–2686.
[2] J. Villen, S.A. Beausoleil, S.A. Gerber, S.P. Gygi, Large-scale phosphorylation analysis of mouse liver, Proc. Natl. Acad. Sci. U.S.A. 104 (2007) 1488–1493.
[3] K. Schmelzle, F.M. White, Phosphoproteomic approaches to elucidate cellular signaling networks, Curr. Opin. Biotechnol. 17 (2006) 406–414.
[4] E. Wakasugi, T. Kobayashi, Y. Tamaki, Y. Nakano, Y. Ito, I. Miyashiro, Y. Komoike, M. Miyazaki, T. Takeda, T. Monden, M. Monden, Analysis of phosphorylation of pRB and its regulatory proteins in breast cancer, J. Clin. Pathol. 50 (1997) 407–412.
[5] L.C. Murphy, G.P. Skliris, B.G. Rowan, M. Al-Dhaheri, C. Williams, C. Penner, S. Troup, S. Begic, M. Parisien, P.H. Watson, The relevance of phosphorylated forms of estrogen receptor in human breast cancer in vivo, J. Steroid Biochem. Mol. Biol. 114 (2009) 90–95.
[6] J. Cicenas, The potential role of Akt phosphorylation in human cancers, Int. J. Biol. Markers 23 (2008) 1–9.
[7] H.J. Lin, F.C. Hsieh, H. Song, J. Lin, Elevated phosphorylation and activation of PDK-1/AKT pathway in human breast cancer, Br. J. Cancer 93 (2005) 1372–1381.
[8] J.Q. Chen, J. Russo, ERalpha-negative and triple negative breast cancer: molecular features and potential therapeutic approaches, Biochim. Biophys. Acta 1796 (2009) 162–175.
[9] C. Grundker, C. Fost, S. Fister, N. Nolte, A.R. Gunthert, G. Emons, Gonadotropin-releasing hormone type II antagonist induces apoptosis in MCF-7 and triple-negative MDA-MB-231 human breast cancer cells in vitro and in vivo, Breast Cancer Res. 12 (2010) R49.
[10] V.G. Keshamouni, R.R. Mattingly, K.B. Reddy, Mechanism of 17-beta-estradiol-induced Erk1/2 activation in breast cancer cells. A role for HER2 AND PKC-delta, J. Biol. Chem. 277 (2002) 22558–22565.
[11] C.K. Osborne, K. Hobbs, J.M. Trent, Biological differences among MCF-7 human breast cancer cell lines from different laboratories, Breast Cancer Res. Treat. 9 (1987) 111–121.
[12] S.E. Burdall, A.M. Hanby, M.R. Lansdown, V. Speirs, Breast cancer cell lines: friend or foe? Breast Cancer Res. 5 (2003) 89–95.
[13] O. Dohr, C. Vogel, J. Abel, Different response of 2, 3,7,8-tetrachlorodibenzo-p-dioxin (TCDD)-sensitive genes in human breast cancer MCF-7 and MDA-MB 231 cells, Arch. Biochem. Biophys. 321 (1995) 405–412.
[14] M. Koutsilieris, C. Reyes-Moreno, I. Choki, A. Sourla, C. Doillon, N. Pavlidis, Chemotherapy cytotoxicity of human MCF-7 and MDA-MB 231 breast cancer cells is altered by osteoblast-derived growth factors, Mol. Med. 5 (1999) 86–97.
[15] E.H. Jo, S.J. Lee, N.S. Ahn, J.S. Park, J.W. Hwang, S.H. Kim, O.I. Aruoma, Y.S. Lee, K.S. Kang, Induction of apoptosis in MCF-7 and MDA-MB-231 breast cancer cells by Oligonol is mediated by Bcl-2 family regulation and MEK/ERK signaling, Eur. J. Cancer Prev. 16 (2007) 342–347.
[16] K. Nohara, F. Wang, S. Spiegel, Glycosphingolipid composition of MDA-MB-231 and MCF-7 human breast cancer cell lines, Breast Cancer Res. Treat. 48 (1998) 149–157.
[17] K. Nesaretnam, R. Stephen, R. Dils, P. Darbre, Tocotrienols inhibit the growth of human breast cancer cells irrespective of estrogen receptor status, Lipids 33 (1998) 461–469.
[18] M. Delehedde, E. Deudon, B. Boilly, H. Hondermarck, Production of sulfated proteoglycans by human breast cancer cell lines: binding to fibroblast growth factor-2, J. Cell. Biochem. 64 (1997) 605–617.
[19] M. Mann, S.E. Ong, M. Gronborg, H. Steen, O.N. Jensen, A. Pandey, Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome, Trends Biotechnol. 20 (2002) 261–268.
[20] M.O. Collins, L. Yu, J.S. Choudhary, Analysis of protein phosphorylation on a proteome-scale, Proteomics 7 (2007) 2751–2768.
[21] H. Matsuda, H. Nakamura, T. Nakajima, New ceramic titania: selective adsorbent for organic phosphates, Anal. Sci. 6 (1990) 911–912.
[22] Y. Ikeguchi, H. Nakamura, Determination of organic phosphates by column-switching high performance anion-exchange chromatography using on-line preconcentration on titania, Anal. Sci. 13 (1997) 479–483.
[23] P.A. Connor, A.J. McQuillan, Phosphate adsorption onto TiO_2 from aqueous solutions: an in situ internal reflection infrared spectroscopic study, Langmuir 15 (1999) 2916–2921.
[24] J. Wu, Q. Shakey, W. Liu, A. Schuller, M.T. Follettie, Global profiling of phosphopeptides by titania affinity enrichment, J. Proteome Res. 6 (2007) 4684–4689.
[25] A.B. Iliuk, V.A. Martin, B.M. Alicie, R.L. Geahlen, W.A. Tao, In-depth analyses of kinase-dependent tyrosine phosphoproteomes based on metal ion-functionalized soluble nanopolymers, Mol. Cell. Proteomics 9 (2010) 2162–2172.
[26] T.E. Thingholm, O.N. Jensen, P.J. Robinson, M.R. Larsen, SIMAC (sequential elution from IMAC), a phosphoproteomics strategy for the rapid separation of monophosphorylated from multiply phosphorylated peptides, Mol. Cell. Proteomics 7 (2008) 661–671.
[27] D.E McNulty, R.S. Annan, Hydrophilic interaction chromatography reduces the complexity of the phosphoproteome and improves global phosphopeptide isolation and detection, Mol. Cell. Proteomics 7 (2008) 971–980.
[28] S.S. Jensen, M.R. Larsen, Evaluation of the impact of some experimental procedures on different phosphopeptide enrichment techniques, Rapid Commun. Mass Spectrom. 21 (2007) 3635–3645.

[29] B. Bodenmiller, L.N. Mueller, M. Mueller, B. Domon, R. Aebersold, Reproducible isolation of distinct, overlapping segments of the phosphoproteome, Nat. Methods 4 (2007) 231–237.
[30] J.K Eng, A.L. McCormack, I. John, R. Yates, An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database, J. Am. Soc. Mass Spectrom. 5 (1994) 976–989.
[31] J. Peng, J.E. Elias, C.C. Thoreen, L.J. Licklider, S.P. Gygi, Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC–MS/MS) for large-scale protein analysis: the yeast proteome, J. Proteome Res. 2 (2003) 43–50.
[32] D. Schwartz, S.P. Gygi, An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets, Nat. Biotechnol. 23 (2005) 1391–1398.
[33] B. Zhang, N.C. VerBerkmoes, M.A. Langston, E. Uberbacher, R.L. Hettich, N.F. Samatova, Detecting differential and correlated protein expression in label-free shotgun proteomics, J. Proteome Res. 5 (2006) 2909–2918.
[34] F. Yates, Contingency tables involving small numbers and the X2 test, J. R. Stat. Soc. (Suppl. 1) (1934) 217–235.
[35] R. Sehgal, A. Berg, J.P. Hegarty, A.A. Kelly, Z. Lin, L.S. Poritz, W.A. Koltun, NOD2/CARD15 mutations correlate with severe pouchitis after ileal pouch-anal anastomosis, Dis. Colon Rectum 53 (2010) 1487–1494.
[36] L.R. Yu, Z. Zhu, K.C. Chan, H.J. Issaq, D.S. Dimitrov, T.D. Veenstra, Improved titanium dioxide enrichment of phosphopeptides from HeLa cells and high confident phosphopeptide identification by cross-validation of MS/MS and MS/MS/MS spectra, J. Proteome Res. 6 (2007) 4150–4162.
[37] Q.R. Li, Z.B. Ning, J.S. Tang, S. Nie, R. Zeng, Effect of peptide-to-TiO_2 beads ratio on phosphopeptide enrichment selectivity, J. Proteome Res. 8 (2009) 5375–5381.
[38] Y. Li, X. Xu, D. Qi, C. Deng, P. Yang, X. Zhang, Novel Fe_3O_4@TiO_2 core-shell microspheres for selective enrichment of phosphopeptides in phosphoproteome analysis, J. Proteome Res. 7 (2008) 2526–2538.
[39] H. Molina, D.M. Horn, N. Tang, S. Mathivanan, A. Pandey, Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry, Proc. Natl. Acad. Sci. U.S.A. 104 (2007) 2199–2204.
[40] B. Thiede, S. Lamer, J. Mattow, F. Siejak, C. Dimmler, T. Rudel, P.R. Jungblut, Analysis of missed cleavage sites, tryptophan oxidation and N-terminal pyroglutamylation after in-gel tryptic digestion, Rapid Commun. Mass Spectrom. 14 (2000) 496–502.
[41] S.C Bendall, C. Hughes, M.H. Stewart, B. Doble, M. Bhatia, G.A. Lajoie, Prevention of amino acid conversion in SILAC experiments with embryonic stem cells, Mol. Cell. Proteomics 7 (2008) 1587–1597.
[42] G.T. Cantin, J.D. Venable, D. Cociorva, J.R. Yates 3rd, Quantitative phosphoproteomic analysis of the tumor necrosis factor pathway, J. Proteome Res. 5 (2006) 127–134.
[43] S.A. Beausoleil, J. Villen, S.A. Gerber, J. Rush, S.P. Gygi, A probability-based approach for high-throughput protein phosphorylation analysis and site localization, Nat. Biotechnol. 24 (2006) 1285–1292.
[44] B. Bodenmiller, J. Malmstrom, B. Gerrits, D. Campbell, H. Lam, A. Schmidt, O. Rinner, L.N. Mueller, P.T. Shannon, P.G. Pedrioli, C. Panse, H.K. Lee, R. Schlapbach, R. Aebersold, PhosphoPep – a phosphoproteome resource for systems biology research in Drosophila Kc167 cells, Mol. Syst. Biol. 3 (2007) 139.
[45] J.V. Olsen, B. Blagoev, F. Gnad, B. Macek, C. Kumar, P. Mortensen, M. Mann, Global, in vivo, and site-specific phosphorylation dynamics in signaling networks, Cell 127 (2006) 635–648.
[46] R.P Zahedi, U. Lewandrowski, J. Wiesner, S. Wortelkamp, J. Moebius, C. Schutz, U. Walter, S. Gambaryan, A. Sickmann, Phosphoproteome of resting human platelets, J. Proteome Res. 7 (2008) 526–534.
[47] F.A. Gonzalez, D.L. Raden, R.J. Davis, Identification of substrate recognition determinants for human ERK1 and ERK2 protein kinases, J. Biol. Chem. 266 (1991) 22159–22163.
[48] X. Yang, D. Gabuzda, Mitogen-activated protein kinase phosphorylates and regulates the HIV-1 Vif protein, J. Biol. Chem. 273 (1998) 29879–29887.
[49] Z. Xiu-mei, L. Bai-lin, S. Min, S. Ji-ye, Expression and significance of ERK protein in human breast carcinoma, Ch. J. Cancer Res. 16 (2004) 269–273.
[50] D.T. Ross, U. Scherf, M.B. Eisen, C.M. Perou, C. Rees, P. Spellman, V. Iyer, S.S. Jeffrey, M. Van de Rijn, M. Waltham, A. Pergamenschikov, J.C. Lee, D. Lashkari, D. Shalon, T.G. Myers, J.N. Weinstein, D. Botstein, P.O. Brown, Systematic variation in gene expression patterns in human cancer cell lines, Nat. Genet. 24 (2000) 227–235.
[51] P.V. Hornbeck, I. Chabra, J.M. Kornhauser, E. Skrzypek, B. Zhang, PhosphoSite: a bioinformatics resource dedicated to physiological protein phosphorylation, Proteomics 4 (2004) 1551–1561.
[52] J.N. Zhou, S. Ljungdahl, M.C. Shoshan, J. Swedenborg, S. Linder, Activation of tissue-factor gene expression in breast carcinoma cells by stimulation of the RAF-ERK signaling pathway, Mol. Carcinog. 21 (1998) 234–243.
[53] H. Chen, G. Zhu, Y. Li, R.N. Padia, Z. Dong, Z.K. Pan, K. Liu, S. Huang, Extracellular signal-regulated kinase signaling pathway regulates breast cancer cell migration by maintaining slug expression, Cancer Res. 69 (2009) 9228–9235.
[54] T. Kinoshita, I. Yoshida, S. Nakae, K. Okita, M. Gouda, M. Matsubara, K. Yokota, H. Ishiguro, T. Tada, Crystal structure of human mono-phosphorylated ERK1 at Tyr204, Biochem. Biophys. Res. Commun. 377 (2008) 1123–1127.
[55] H.N. Wong, M.A. Ward, A.W. Bell, E. Chevet, S. Bains, W.P. Blackstock, R. Solari, D.Y. Thomas, J.J. Bergeron, Conserved in vivo phosphorylation of calnexin at casein kinase II sites as well as a protein kinase C/proline-directed kinase site, J. Biol. Chem. 273 (1998) 17227–17235.
[56] E. Chevet, J. Smirle, P.H. Cameron, D.Y. Thomas, J.J. Bergeron, Calnexin phosphorylation: linking cytoplasmic signalling to endoplasmic reticulum lumenal functions, Semin. Cell Dev. Biol. 21 (2010) 486–490.
[57] C.J. Chen, J.A. Traugh, Expression of recombinant elongation factor 1 beta from rabbit in *Escherichia coli*. Phosphorylation by casein kinase II, Biochim. Biophys. Acta 1264 (1995) 303–311.
[58] P. Jowsey, N.A. Morrice, C.J. Hastie, H. McLauchlan, R. Toth, J. Rouse, Characterisation of the sites of DNA damage-induced 53BP1 phosphorylation catalysed by ATM and ATR, DNA Repair (Amst) 6 (2007) 1536–1544.
[59] M.J. Lim, X.W. Wang, Nucleophosmin and human cancer, Cancer Detect. Prev. 30 (2006) 481–490.
[60] W.J. Qian, M.B. Goshe, D.G. Camp 2nd, L.R. Yu, K. Tang, R.D. Tang, Phosphoprotein isotope-coded solid-phase tag approach for enrichment and quantitative analysis of phosphopeptides from complex mixtures, Anal. Chem. 75 (2003) 5441–5450.